Use of One-Point Coverage Representations, Product Space Conditional Event Algebra, and Second-Order Probability Theory for Constructing and Using Probability-Compatible Inference Rules in Data-Fusion Problems

I. R. Goodman
SSC San Diego

ABSTRACT

This paper covers issues relating to the establishment of a sound and conditional probability-compatible rationale for generating linguistic-based inference rules concerning a population. By extending previous preliminary results, we detail, in a fully rigorous manner and within the confines of traditional probability theory, that a comprehensive technique can be derived that converts linguistic-based conditional information, couched only in fuzzy-logic terms, into naturally corresponding conditional probabilities. In turn, we demonstrate how such typically underconstrained conditional probabilities can be combined for suitable conclusions and decision-making, via a new use of second-order probability logic. This research is part of the ongoing SSC San Diego In-house Laboratory Independent Research FY 01 project CRANOF (a Complexity-Reducing Algorithm for Near-Optimal Fusion).

INTRODUCTION

Programmatics

This paper documents one aspect of the ongoing FY 01 In-house Laboratory Independent Research Project CRANOF (a Complexity-Reducing Algorithm for Near-Optimal Fusion), Project ZU014, with Principal Investigator, Dr. D. Bamber, and co-investigator, Dr. I. R. Goodman (both SSC San Diego), and with associate support from Dr. W. C. Torrez (SSC San Diego) and Prof. H. T. Nguyen (Department of Mathematical Sciences, New Mexico State University and U.S. Navy American Society for Engineering Education Fellow during summers at SSC San Diego). A preliminary version of this paper can be found in [1, section 3.3].

Background on Underconstrained Conditional Probability Problems

Philosophy of Approach and General Motivations

To improve the timeliness and accuracy of decision-supported human decision-making, one is faced with an array of crucial problems, including how to handle large amounts of incoming and uncertain information from disparate sources. These sources can be human-based or mechanical-based, and the information can arrive in different forms, such as qualitative and linguistic, numerical and statistical-probabilistic, or some mixture of both. At SSC San Diego, the CRANOF project addresses such crucial issues solely within the realm of statistics and probability. The issue of underconstrained or underspecified probabilities is treated by a novel use of second-order probabilities (i.e., probabilities of probabilities) in Bayesian framework. Underconstrained probabilities arise in a wide variety of problems, including quantitatively formulated rule-based systems, tracking and correlation, assessment of network intrusions, information retrieval, and simulation of human behavior in war games. This paper serves as a beginning extension of the capabilities of CRANOF to include linguistic-based information.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **AUG 2001**	2. REPORT TYPE	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **Use of One-Point Coverage Representations, Product Space Conditional Event Algebra, and Second-Order Probability Theoryfor Constructing and Using Probability-Compatible Inference Rules in Data-Fusion Problems**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Space and Naval Warfare Systems Center,53560 Hull Street,San Diego,CA,92152-5001**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S) 11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This paper covers issues relating to the establishment of a sound and conditional probability-compatible rationale for generating linguistic-based inference rules concerning a population. By extending previous preliminary results, we detail, in a fully rigorous manner and within the confines of traditional probability theory, that a comprehensive technique can be derived that converts linguistic-based conditional information, couched only in fuzzy-logic terms, into naturally corresponding conditional probabilities. In turn, we demonstrate how such typically underconstrained conditional probabilities can be combined for suitable conclusions and decision-making, via a new use of second-order probability logic. This research is part of the ongoing SSC San Diego In-house Laboratory Independent Research FY 01 project CRANOF (a Complexity- Reducing Algorithm for Near-Optimal Fusion).

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**12**	

Quantitatively Formulated Rule-Based Systems

Consider quantitatively formulated rule-based systems, with the rules or conditional relations symbolized typically as $(a_1 \mid b_1)$, $(a_2 \mid b_2)$,...—read "if b_1, then a_1" (or equivalently, "a_1, given b_1," etc.), "if b_2, then a_2,"..., where events or sets a_1, b_1, a_2, b_2,... may themselves represent quite complicated logical combinations of simpler events or sets, and where it may or may not be known what logical relations exist among such events. Each such rule is also assigned quantitative reliability in the form of naturally corresponding conditional probabilities. Thus, for some otherwise unspecified probability measure P, rule $(a|b)$ is assigned value $P(a|b) = P(ab)/P(b)$, the conditional probability of a given b, using standard Boolean and probability notation and assuming antecedent probability $P(b) > 0$. Because typical rule $(a|b)$ is not perfect, in general $P(a|b) < 1$, but, on the other hand, one would expect $P(a|b)$ to be reasonably high. A common problem that such rule-based systems address is: Consider incoming information in the form of events, d_1,..., d_n, possibly gleaned from different sources, such as d_1 = "visibility is up to 1 mile," d_2 = "winds between 15 mph and 30 mph," d_3 = "enemy movement detected last night in Sector C,"..., d_n = "political situation with enemy country Q at level R," and a collection of reasonably related rules, such as $(a_1|b_1)$, $(a_2|b_2)$,..., $(a_m|b_m)$, where the a_j, b_j involve not only parts or all of the d_j (or various logical combinations of them), but possibly other related events (or logical combinations of such). Then, one wishes to test for viability of possible decisions, based upon this information, such as c_1 = "fully successful attack by us can be accomplished by attacking in Sectors C or D," c_2 = "partially successful attack by us can be accomplished by attacking Sectors D or H,"... . Symbolically, one is considering the validity or degree of validity of the *entailment schemes* $G_i = [(a_1|b_1),..., (a_n|b_n); (c_i|d)]$, i = 1, 2,... , where $d = d_1\&...\&d_n$ (conjunction of all data), and where $((a_1|b_1),..., (a_n|b_n))$ can be considered the *premise set* of G_i and $(c_i|d)$ its *potential conclusion*. Ideally, one would like to know just what each $P(c_i|d)$ would be, based on having either, say, the *exact threshold situation* holding, i.e., $P(a_j|b_j) = t_j$, j = 1,..., n, or, the *lower bound threshold* situation holding, i.e., having $P(a_j|b_j) \geq t_j$, where all the thresholds t_j are known or estimable in either situation. However, in general, it is readily demonstrated that the n equalities (or inequalities) are not enough to determine P and/or $P(c_i|d)$ completely. Thus, one is faced with the problem of best estimating, in some sense, just what P and/or $P(c_i|d)$ should be.

Adams' Approach to Analyzing Quantitatively Formulated Rule-Based Systems

In a series of papers [2, 3], E. W. Adams proposed, in effect, the estimate of $P(c_j|d)$ to be a pessimistic one in the form of his "minimum conclusion" function, using multivariable abbreviation t_J for $(t_j)_{j \text{ in } J}$, $(a|b)_J$ for $(a_j|b_j)_{j \text{ in } J}$, $P(a|b)_J \geq t_J$ for $P(a_j|b_j) \geq t_j$, j in J, 1_J for column vector of all 1's indexed by J, etc.,

$$\text{estimate}_{HPL} \text{ of } (P(c_i|d) \text{ from } G_i)$$
$$= \text{minconc}(G_i)(t_J) = \inf\{P(c_i|d): \text{for all possible probability measures P such that } P(a|b)_J \geq t_J\}, \quad (1)$$

with $P(c_i|d)$ for the exact threshold situation analogously estimated. The subscript $(\ )_{HPL}$ is used to indicate "High Probability Logic," since Adams also introduced the idea of an entailment scheme being *HP-valid* or *HP-invalid*, which, in the case of any G_i here simply means for the former that

$$G_i \text{ is HPL-valid} \quad \text{iff} \quad \underset{(t_J \uparrow 1_J)}{\text{limit}}(\text{minconc}(G_i)(t_J)) = 1. \tag{2}$$

But, unfortunately, both the minconc function and its limiting forms to test for HPL-validity/invalidity produce a number of results very much at odds with commonsense reasoning, including the fact that three very fundamental entailment schemes, *transitivity* (or *hypothetical syllogism*) [(a|b), (b|c); (a|c)] (the heart of any rule-based system); *contraposition* [(a|b); (b′|a′)]; and *strengthening of antecedent* [(a|b); (a|bc)] are all HPL-invalid. In fact, one can find P’s that satisfy their premise thresholds for any choice of t_J close to (but not exactly equal to) 1_J, but for which the corresponding conclusion probabilities are arbitrarily close to (or actually equal to) 0. Moreover, more generally, Eq. (2) can be complemented by the fact that any

$$G_i \text{ is HPL-invalid} \quad \text{iff} \quad \underset{(t_J \uparrow 1_J)}{\text{limit}}(\text{minconc}(G_i)(t_J)) = 0. \tag{3}$$

Finally, Adams pointed out another type of validity, CPL (Certainty Probability Logic), that, although still based on the minconc function, can be characterized as "too optimistic" in contrast with HPL, whereby the criterion is

$$G_i \text{ is CPL-valid} \quad \text{iff} \quad \text{minconc}(G_i)(1_J)) = 1. \tag{4}$$

Close connections exist between CPL validity/invalidity (the latter satisfying a relation analogous to that of Eq. (3)) and that of CL (classical logic) validity or invalidity, noting

$$G_i \text{ is CL-valid} \quad \text{iff} \quad \&(b' \vee ab)_J \leqslant d' \vee c_i d. \tag{5}$$

(For further analysis, criticism, and extension of Adams’ ideas, see [3].)

CRANOF Approach to Analyzing Quantitative Rule-Based Systems and Other Underconstrained Probability Problems

The previous conclusions show that the minconc function is not a reasonable measure (for reasonably high thresholds) of the degree of validity/invalidity of an entailment scheme and also show that the HP-validity/invalidity test is too stringent. Therefore, it seemed natural to replace the extremal minconc function by the more moderating *meanconc* function (well-justified from decision analysis in the form of conditional expectation and justified as always admissible, least-squares error, etc.—see any standard texts such as Rao [4] or Wilks [5]) within a Bayesian framework, where the unknown probability measure P here is treated as a random quantity with some appropriately assigned prior distribution, subject to the given premise set threshold constraints. Utilizing additional new theoretical results [6], an "optimal" choice of prior or priors essentially must come from the well-known Dirichlet family of distributions. It should be noted that, unlike the minconc function, the meanconc function in the unity-limiting threshold case can take on nontrivial values and, in a natural sense, at any fixed threshold level, provides a reasonable measure of degree of validity of that entailment scheme under consideration. In particular, in full agreement with commonsense reasoning, transitivity, contraposition, and strengthening of antecedent are all SOPL-valid, where SOPL stands for Second-Order Probability Logic and where one defines validity of any G_j as

$$G_i \text{ is SOPL-valid} \quad \text{iff} \quad \lim_{(t_J \uparrow 1_J)}(\text{meanconc}(G_i)(t_J)) = 1, \tag{6}$$

SOPL-validity depending on some degree, of course, on the particular choice of prior for P. However, it has been pointed out (Bamber [7] and personal communications) that the limit in Eq. (4) remains the same as if the prior of P is a uniform distributional one, when the corresponding probability density function is bounded uniformly above and below (from zero) over its *natural* domain (again, see references).

Also, see [8] for additional background on both the theoretical structure of the meanconc function and its practical implementational form CRANOF—whereby a significant reduction in the complexity of computing meanconc(G_i)(t_J) is achieved by, in effect, reducing the premise set of G_i to a single constraint, also taking into account the unity-limiting threshold behavior of meanconc ([7]). Finally, Table 1 is presented below to illustrate a few typical evaluations of meanconc(G) for relatively simple entailment schemes G with P assigned a uniform prior distribution [8].

TABLE 1. Abridged table of calculations of degree-of-entailment functions, minconc and meanconc, for fixed threshold levels, and a comparison of CPL-, SOPL-, and HPL-validities for different types of entailment schemes.

Name of Entailment Scheme $D = [(a\|b)_J; (c\|d)]$	Given Levels of Premises: $P(a\|b)_J = t_J$, for otherwise arbitrary prob. meas. P	minconc(D)(t_J) (inequality threshold form)	meanconc(D)(t_J), assuming uniform prior for P's (exact threshold form)	D is CPL-valid?	D is SOPL-valid?	D is HPL-valid?
Cautious Monotonicity: [(a|b),(c|b); (a|bc)]	P(a|b) = s, P(c|b) = t	$\geq \max(s+t-1,0)$	$\geq \max(s+t-1,0)$	YES	YES	YES
Transitivity: [(a|b), (b|c); (a|c)]	P(a|b) = s, P(b|c) = t	0	$= st + (1-t)/2 - p(s,t)/q(s,t)$, $p(s,t) = s(1-s)(2s-1)t(1-t^2)$, $q(s,t) = t+2t^2+ (s(1-s)(1-t)(2+3t-t^2)$	YES	YES	NO
Contraposition: [(a|b); (b′|a′)]	P(a|b) = t	0	$1/t + \frac{(1-t)\log(1-t)}{t^2}$	YES	YES	NO
Positive Conjunction: [(a|b),(a|c); (a|bc)]	P(a|b) = t, P(a|c) = t	0	$(1+t)/3 + [((1+t)(2-t)/(3t))\,\theta(t)]$, $\theta(t) = (t^2/4)[\log((2-t)/t)]/(1-t) - ((1-t)^2/4)\cdot\log((1+t)/(1-t))$	YES	YES	NO
Nixon Diamond: [(ab|c),(d|a),(d′|b); (d|c)]	P(ab|c) = s, P(d|a) = t, P(d′|b) = t	0	1/2	YES	NO	NO
Abduction: [(a|b), a; b]	P(a|b) = s, P(a) = t	0	If $s \geq t$: $t/(2s)$, If $s < t$: $\frac{t^3\, s(1-t)^2}{2(t^2-2st+s)^2}$	NO	NO	NO

EXTENDING APPLICABILITY OF CRANOF TO LINGUISTIC-BASED SYSTEMS

In considering linguistic-based information in rule-based systems and in formulating the linguistic analogue of the underconstrained conditional (including unconditional) probability problem, the role of fuzzy logic comes immediately to mind. This is based in part on the great practical success of fuzzy logic in running systems such as elevators, washing machines, etc., and on the now very large body of scientific literature supporting the modeling of linguistic information, relations, and decision processes via fuzzy logic. (See, e.g., past *Proceedings of IEEE International Conferences on Fuzzy Systems* or the *Proceedings of the Joint Conference on Information Sciences*, as well as basic texts, such as Dubois & Prade's now classic treatise [9] and Nguyen & Walker's [10].)

On the other hand, there still exists a lively controversy considering the merits of using probability theory and techniques in place of fuzzy logic and vice versa. (See Goodman's summary and listing of literature papers directly involved in this controversy [11].) This leads to the following area in which this author and H. T. Nguyen have played some role over the past several years: *the issue of the possible direct connection between fuzzy logic and probability theory* [12, 13, and 1]. Until this is completely resolved, it is this author's opinion that a comprehensive view of data fusion, which both theoretically and practically integrates linguistic-based information with probabilistic-based information, will not be achieved. In particular, this applies to rule-based systems, where the fuzzy logic community has developed a common approach that is claimed to be more satisfactory than any probability approach.

This paper once again points out the existence of deep, but tractable, relations among fuzzy logic, linguistic-based principles, probability theory, and commonsense reasoning mainly through the use of two basic mathematical tools: SOPL/CRANOF (as briefly described in the first section), and the representation theory of fuzzy sets by the one-point coverages of random sets (see [12, 13]) in conjunction with other recently developed mathematical tools (*conditional and relational event algebra* [14; 15, section 3]). In particular, homomorphic-like relations were established, connecting fuzzy-logic concepts and corresponding random-set concepts, where each fuzzy-set membership function involved is, in effect, interpreted as the weakest way to specify any of a class of corresponding random subsets of the fuzzy set's domain. These relations include natural random-set interpretations of various combinations of fuzzy-logic operators and Zadeh's well-known "extension theorem." This time, these connections are extended to include the formulation and use of inference rules obtained from a population of interest. The results presented here extend preliminary efforts provided in Goodman & Nguyen [1], where it was demonstrated that one type of fuzzy-logic approach to the modeling of inference rules for a population, relative to a given collection of attributes, using the ratio of fuzzy cardinalities or averaged membership level of the attributes, could also be interpreted in a probability framework. In addition, by using similar techniques, it is shown how other fuzzy-logic concepts, commonly thought of as not directly relating to probability, may now also be put into a complete probabilistic setting, including the illustration for normalization of membership functions.

MATHEMATICAL RESULTS ESTABLISHING GENERAL FUZZY LOGIC POPULATION CONDITIONING PROBLEM AS AN UNDERCONSTRAINED CONDITIONAL PROBABILITY PROBLEM TREATABLE VIA SOPL/CRANOF

As in the previous sections, standard Boolean algebra and probability theory notation will be employed, with [0,1] indicating unit interval; {0,1} indicating the two element set containing 0, 1; **R** indicating the real (or Euclidean) line and $\mathbf{R}^m$ indicating the real (or Euclidean m-space), P(D) indicating the power class of D (sometimes written 2^D—the class of all subsets of D), etc. "Equal by definition" is denoted as $=_d$. For background on copulas, see Schweizer & Sklar [16] and the recent excellent monograph by Nelsen [17]. Recall that copulas are any joint cdf's (cumulative probability distribution functions), all of whose one-dimensional marginal cdf's correspond to identical uniformly distributed rv's (random variables) over [0,1].

Theorem 1. Modification of Goodman [18]

Let D be a finite set, f, g:D→[0,1] any two fuzzy-set membership functions, and cop: $[0,1]^{D\times D}\to[0,1]$ any copula with that domain, with (x,y)-marginal copulas indicated by, e.g., $cop_{x,y}$, x, y in D, etc. Then:

(i) There is a probability space ($\mathbf{\Omega}$,B,P) and a joint collection of 0-1-valued rv's, $Z_{f,x}$, $Z_{g,y}$:$\mathbf{\Omega}\to\{0,1\}$, for all x, y in D with overall joint cdf $F_{f,g,cop} = cop_o((F_{f,x})_{x \text{ in } D}, (F_{g,y})_{y \text{ in } D})$: $\mathbf{R}^{D\times D}\to[0,1]$ (via Sklar's Theorem [16]), and, indicating the joint marginal (x,y)-components of cop, as $cop_{x,y}$, the joint cdf of ($Z_{f,x}$, $Z_{g,y}$) is, correspondingly, $F_{f,g,cop,x,y}(.,..) = cop_{x,y}\circ(F_{f,x}(.), F_{g,y}(..))$, where ^{o} indicates functional composition and $F_{f,x}$, $F_{g,y}$ are each one-dimensional cdf's corresponding to mass-point probability functions $h_{f,x}$, $h_{g,y}$, respectively, where

$$\begin{aligned}P(Z_{f,x} = 1) &= h_{f,x}(1) = f(x);\ P(Z_{f,x} = 0) = h_{f,x}(0) = 1-f(x);\\ P(Z_{g,y} = 1) &= h_{g,y}(1) = g(y);\ P(Z_{g,y} = 0) = h_{g,y}(0) = 1-g(y);\end{aligned} \tag{7}$$

whence

$$F_{f,x}(s)=\begin{cases}0, \text{ if } s < 0,\\ 1-f(x), \text{ if } 0 \leqslant s < 1,\\ 1, \text{ if } 1 \leqslant s;\end{cases}\quad F_{g,y}(s)=\begin{cases}0, \text{ if } s < 0,\\ 1-g(y), \text{ if } 0 \leqslant s <1,\\ 1, \text{ if } 1 \leqslant s;\end{cases}\ \text{all } x, y \text{ in } D \tag{8}$$

(ii) Define random sets S(f, cop), S(g, cop):$\mathbf{\Omega}\to P$(D), S(f, g, cop): $\mathbf{\Omega}\to P(D)\times P(D)$ as follows, for each ω in $\mathbf{\Omega}$:

$$\begin{aligned}&S(f, g, cop)(\omega) = S(f, cop)(\omega)\times S(g, cop)(\omega) = \{(x,y): x, y \text{ in } D, Z_{f,x}(\omega)\, Z_{g,y}(\omega) = 1\};\\ &S(f, cop)(\omega) = \{x: x \text{ in } D, Z_{f,x}(\omega) = 1\};\quad S(g, cop)(\omega) = \{y: y \text{ in } D, Z_{g,y}(\omega) = 1\};\end{aligned} \tag{9}$$

whence, by straightforward combinatoric considerations, the entire probability distributions of the marginal random subsets of D, S(f, cop), S(g, cop), as well as the joint random subset of D×D, S(f, g, cop), are completely determined.

(iii) For any x, y in D, the following equality of one-point coverage events hold:

$$(x \text{ in } S(f, \text{cop})) = (Z_{f,x} = 1) \;;\; (y \text{ in } S(g, \text{cop})) = (Z_{g,y} = 1); \tag{10}$$

$$((x,y) \text{ in } S(f, g, \text{cop})) = (x \text{ in } S(f, \text{cop})) \;\&\; (y \text{ in } S(g, \text{cop})) = (Zf,x = 1) \;\&\; (Z_{g,y} = 1). \tag{11}$$

(iv) For any x, y in D, the following *one-point coverage representations* for f, g hold:

$$P(x \text{ in } S(f, \text{cop})) = P(Z_{f,x} = 1) = f(x) \;;\; P(y \text{ in } S(g, \text{cop})) = P(Z_{g,y} = 1) = g(y); \tag{12}$$

$$\begin{aligned} P((x \text{ in } S(f, \text{cop})) \;\&\; (y \text{ in } S(g, \text{cop}))) &= P((Z_{f,x} = 1) \;\&\; (Z_{g,y} = 1)) \\ &= 1 - P(Z_{f,x} = 0) - P(Z_{g,y} = 0) + P((Z_{f,x} = 0) \;\&\; (Z_{g,y} = 0)) \\ &= 1 - P(Z_{f,x} = 0) - P(Z_{g,y} = 0) + P((Z_{f,x} \leq 0) \;\&\; (Z_{g,y} \leq 0)) \\ &= 1 - (1-f(x)) - (1-g(y)) + F_{f,g,\text{cop}_{x,y}}(0, 0) \\ &= f(x) + g(y) - 1 + \text{cop}_{x,y}(1-f(x), 1-g(y)) \\ &= f(x) + g(y) - \text{cocop}_{x,y}(f(x), g(y)) \\ &=_d \text{cop}_{x,y}\hat{}(f(x), g(y)), \end{aligned} \tag{13}$$

where we use the relation

$$\begin{aligned} F_{f,g,\text{cop}_{x,y}}(0, 0) &= \text{cop}_{x,y}{}^{\circ}(F_{f,x}(0), F_{g,y}(0)) \\ &= \text{cop}_{x,y}{}^{\circ}(h_{f,x}(0), h_{g,y}(0)) \\ &= \text{cop}_{x,y}{}^{\circ}(1-f(x), 1-g(y)) \end{aligned}$$

and where the functions cocop, cop^ are called the *cocopula, survival copula*, respectively, of cop (the latter apparently being the special designation of Nelsen for modular transform [17, section 2.6]), where, for any s, t in [0,1]:

$$\text{cocop}(s, t) =_d 1 - \text{cop}(1-s, 1-t) \;;\; \text{cop}\hat{}(s, t) =_d s+t - \text{cocop}(s,t). \tag{14}$$

(v) Specializing (iv) for x = y in D arbitrary,

$$P(x \text{ in } S(f, \text{cop}) \cap S(g, \text{cop})) = P((x,x) \text{ in } S(f, g, \text{cop})) = \text{cop}_{x,y}\hat{}(f(x), g(x)). \tag{15}$$

(vi) As copula cop is allowed to vary arbitrarily, the full solution set of distribution-distinct random subsets of D that are one-point coverage equivalent to f, g, respectively in the sense of Eq. (12), is exhausted. ■

Remark 1. Note first that cocop is the DeMorgan transform of cop—so that if one thinks of cop as a generalized conjunction or "and" operator—as in fuzzy logic (with the usual desirable properties of being nondecreasing in its arguments and having appropriate boundary properties when one of the arguments is 0 or 1), then, naturally, cocop can be thought of as a general disjunction or "or" operator. Nelsen [17, section 2], shows that the survival copula is always a legitimate copula and shows the characterization

$$\text{cop}\hat{} = \text{cop} \quad \text{iff} \quad \text{cop is } \textit{radially} \text{ symmetric}, \tag{16}$$

where the latter means that the joint r.v. Y represented by cop is such that $Y - (1/2, 1/2)$ and $(1/2, 1/2) - Y$ have the same distribution. In particular, radial symmetry—and hence the validity of Eq. (16)—holds for all Gaussian copulas Ψ_ρ (Ψ-1(.), Ψ-1(..)), where Ψ_ρ is the joint cdf of distribution Gaussian $(0_2, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix})$ and Ψ is the cdf of the standardized one-dimensional Gaussian distribution Gaussian (0,1) and all of Frank's Archimedean copula family [17], [19] (i.e., associative, commutative with cop(s,s) < s, for 0 < s < 1)—which includes the copulas prod and minsum, as well as the special copula min, where for all s, t in [0,1], min, prod are the usual arithmetic minimum and product of s, t, respectively, while minsum(s,t) is given as

$$\mathrm{minsum}(s,t) = \min(s+t-1,\ 0). \tag{17}$$

■

Theorem 2. Extension of Goodman & Nguyen [13]

Suppose that D is a finite set, f, g:D→[0,1] are any two fuzzy set membership functions, cop: $[0,1]^{D\times D}\to[0,1]$ is any copula with that domain, and w:D→[0,1] is a probability function. Define

$$((f|g)_{cop,w} =_d \sum_{x\in D} (\ w(x)\cdot cop\hat{}(f(x), g(x))) \ /\ \sum_{x\in D} (\ w(x)\cdot g(x)). \tag{18}$$

Then, in the sense of Theorem 1, there is a probability space (Ω,B,P) and random sets S(f, cop), S(g, cop): Ω →*P*(D), S(f,g, cop): with the one-point coverage relations holding as in Eqs. (12), and, without loss of generality, there exists a random variable V:Ω→D, independent of S(f, g, cop), and hence of S(f, cop), S(g, cop), such that the probability function of V is w, so that

$$(f|g)_{cop,w} = P(a_{f,\ cop} \mid b_{g,\ cop}), \tag{19}$$

an ordinary conditional probability, where events $a_{f,cop}$, $b_{g,cop}$ in B are defined as the two-stage randomization events

$$a_{f,\ cop} =_d (V \text{ in } S(f,\ cop))\ ,\quad b_{g,\ cop} = (V \text{ in } S(g\ cop))\ , \tag{20}$$

so that in reduced form,

$$P(a_{f,\ cop} \mid b_{g,\ cop}) = P(a_{f,\ cop} \& b_{g,\ cop} \mid b_{g,\ cop}) = P(V \text{ in } S(f,\ cop) \cap S(g,\ cop)) \ /\ P(V \text{ in } S(g\ cop)). \tag{21}$$

Proof: Use the usual conditioning property of probabilities, independence of V, and Eq. (11) at each outcome of r.v. V,

$$\begin{aligned} P(V \text{ in } S(f,cop) \text{ and } V \text{ in } S(g,\ cop)) &= E_V(P(V \text{ in } S(f,cop) \text{ and } V \text{ in } S(g,\ cop) \mid V)) \\ &= E_V(cop\hat{}(f(V),\ g(V))) = \sum_{x\in D} (\ w(x)\cdot cop\hat{}(f(x),\ g(x))). \end{aligned} \tag{22}$$

Similarly (and more simply), now using Eq. (12) in place of Eq. (13),

$$P(V \text{ in } S(g,\ cop)) = E_v(P(V \text{ in } S(g,cop) \mid V)) = E_v(g(V)) = \sum_{x\in D} (\ w(x)\cdot g(x)). \tag{23}$$

The desired results hold by dividing Eq. (22) by Eq. (23). ■

Remark 2 and an Example. In Theorem 2, for the special case of w corresponding to a uniform distribution over *population* D, canceling the 1/card(D) factor, and usually—but not always choosing cop to be either min or prod—the numerator of the quantity $(f|g)_{cop,\ w}$ reduces to the popular fuzzy-logic concept of the *fuzzy cardinality* of f "and" g for population D, i.e., to what extent the entire population D has characteristics described by f "and" g, while, similarly, the denominator represents the fuzzy cardinality of g (by itself) for population D. In turn, the arithmetic division of these, i.e., the quantity $(f|g)_{cop,w}$ becomes the *relative fuzzy cardinality* of f "and" g for D *compared to fuzzy cardinality* of g for D, i.e., the overall fuzzy conditioning of f to g with respect to population D. The latter, beginning with Zadeh's ideas [20, 21], followed by Dubois & Prade's modifications [22], and Kosko's related concept of *fuzzy subsethood* [23], are used ubiquitously in the fuzzy-logic community for reasoning. In this process, one considers the premise set of a particular linguistic entailment of interest, the latter being formally the same as the probability-framed previous $G_i = [(a|b)_J; (c_i|d)]$, but now where each $(a_j|b_j)$ is replaced by a fuzzy conditional—in its general form *the same* as $(f_j|g_j)_{cop,w}$ —formed as in Eq. (18), now with f replaced by f_j, g by g_j (for possibly pre-logically compounded fuzzy-set membership functions), j in J; and with similar remarks applicable to the potential conclusion $(c_i|d)$ replaced by $(f_{o,i}|g_o)_{cop,w}$, for some fuzzy sets $f_{o,i}$, g_o, etc. But, Theorem 2 (with suitable modifications, where required) essentially shows that any such $(f_j|g_j)_{cop,w} = P(a_{fj,\,cop} \mid b_{gj,cop})$, with a similar relation holding the potential conclusion. Moreover, the variability of P subject to whatever arbitrary but fixed levels t_J are set for the premise collection holds in the same meaningful manner as in the case where one began the problem in a probability framework, i.e., for typical entailment schemes of the form G_i. As an application of this, suppose one considers the transitivity scheme, which Zadeh has also considered and modeled his premise set as indicated above, but has used a method solely developed within fuzzy logic for determining what the appropriate conclusion should be [21]. Thus, three attributes are present, where, e.g., population D here is the set of all enemy ships in area A, "ships with type 1 weapons onboard" corresponds to known or estimated fuzzy-set membership function f over D; "ships with elongated hulls" corresponds to known or estimated fuzzy-set membership function g over D; "ships with signature pattern Q" corresponding to known or estimated fuzzy-set membership function h over D. Moreover, other truth modifiers may be present, such as "it is mostly true," "it is somewhat true," etc. Here, for simplicity, suppose for the premise set, one actually has "it is highly true that the enemy ships in A with signature pattern Q have elongated hulls," "it is moderately likely that an enemy ship in A with an elongated hull has type 1 weapons onboard." Can one conclude "it is x-likely that an enemy ship in A with signature pattern Q has type 1 weapons onboard," where the degree of truth x is to be determined? Assume that "it is highly true" is represented by a known or estimated fuzzy-set membership function M over [0,1], which is monotone increasing, "it is moderately likely" is also represented by a (different—not as steep toward 1 as M, etc.) known or estimated fuzzy-set membership function N over [0,1], where M(r) = N(r) = r, for r = 0 or 1. Hence, for any arbitrary levels s, t in [0,1], the conditional fuzzy relations here are, for some choice of copula and population weighting function w,

$$\begin{aligned} M((f|g)_{cop,w}) = s\ ,\ N((g|h)_{cop,w}) = t \ \text{ iff } \ & (f|g)_{cop,w} = M^{-1}(s)\ ,\ (g|h)_{cop,w} = N^{-1}(t) \\ \text{iff, using Theorem 2, } \ & P(a_{f,\,cop} \mid b_{g,cop}) = M^{-1}(s), \\ & P(b_{g,\,cop} \mid c_{h,cop}) = N^{-1}(t). \end{aligned} \tag{24}$$

Thus, for any given levels s, t, one can now consider the SOPL-estimate of the potential conclusion for transitivity, $P(a_{f,\ cop} \mid b_{g,cop})$, with respect to the premise set above at thresholds s, t, where the entire entailment scheme is

$$G = [(a_{f,\ cop} \mid b_{g,cop}), (b_{g,\ cop} \mid c_{h,cop}); (a_{f,\ cop} \mid c_{h,cop})]; \quad (25)$$

$$\text{meanconc}(G)(M^{-1}(s), N^{-1}(t)) = E_P(P(a_{f,\ cop} \mid c_{h,cop}) \mid$$
$$P(a_{f,\ cop} \mid b_{g,cop}) = M^{-1}(s)\ ,\ P(b_{g,\ cop} \mid c_{h,cop}) = N^{-1}(t)). \quad (26)$$

In turn, Table 1 shows that under a uniform distributional assumption on what P could be, subject to its constraints in the premise set of G, for any given s, t in $[1/2, 1]$

$$\text{meanconc}(G)(M^{-1}(s), N^{-1}(t)) = \rho(M^{-1}(s), N^{-1}(t)),$$

where, for any s, t in $[1/2, 1]$,

$$\rho(s,t) =_d st + (1-t)/2 - p(s,t)/q(s,t);\ \ p(s,t) =_d s(1-s)(2s-1)t(1-t^2);$$
$$q(s,t) =_d t+2t^2 + (s(1-s)(1-t)(2+3t-t^2)), \quad (27)$$

where,

$$\rho(s,t) \approx \rho_o(s,t) =_d st + (1-t)/2\ ,\ \text{for values of s, t sufficiently close to 1.} \quad (28)$$

Hence, the posterior conditional (given the premise constraints for any s, t) is approximately equal to $\rho_o(M^{-1}(s), N^{-1}(t))$, which can be interpreted also as a truth modifier with respect to two variables, noting its limit is unity as s, t approach unity, etc. Of course, all of the above applies to any fuzzy-logic entailment scheme relative to the original premise sets utilizing overall fuzzy conditioning for some population D.

Remark 3. In the same spirit of Theorem 2, other fuzzy-logic concepts can now be fully interpreted. Due to space limitations, only the example of fuzzy normalization will be considered here. In this situation, a fuzzy membership function, say, $f:D\to[0,1]$ is given, followed by its normalization function $\text{norm}(f):D\to[0,1]$, which is now obviously a legitimate probability function over finite population D, where

$$\text{norm}(f) = \Big(1/\sum_{x\in D}(f(x))\Big)\cdot f\ . \quad (29)$$

But, if one considers, à la Theorem 1, for any choice of copula cop, a probability space (Ω,B,P), for which, without loss of generality, there is both a random set $S(f,\ cop):\Omega\to P(D)$ and an independent random variable $V:\Omega\to D$ uniformly distributed over D, with the one-point coverage relation holding

$$P(x \text{ in } S(f,\ cop)) = f(x),\ \text{all } x \text{ in } D, \quad (30)$$

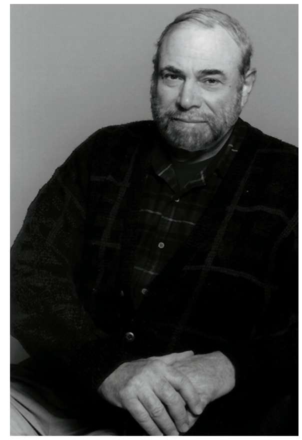
for any x in D, specializing Eq. (23) with g replaced by f,

$$P(V = x \mid V \text{ in } S(f,\ cop)) = P(V = x \text{ and } x \text{ in } S(f,\ cop))\ /\ P(V \text{ in } S(f,\ cop))$$
$$= ((1/\text{card}(D))\cdot f(x))\ /\sum_{x\in D}(1/\text{card}(D))\cdot g(x)\Big) = \text{norm}(f)(x), \quad (31)$$

showing fuzzy normalization is actually a simple conditional probability restriction of the two-stage randomization for one-point coverages. A future paper will deal with related issues.

REFERENCES

1. Goodman, I. R. and H. T. Nguyen. 1999. "Application of Conditional and Relational Event Algebra to the Defining of Fuzzy Logic Concepts," *Proceedings of Signal Processing, Sensor Fusion & Target Recognition VIII*, Society of Photo-Optical Instrumentation Engineers (SPIE), vol. 3720, pp. 37–46.
2. Adams, E. W. 1986. "On the Logic of High Probability," *Journal of Philosophical Logic*, vol. 15, pp. 255–279.
3. Adams, E. W. 1996. "Four Probability-Preserving Properties of Inferences," *Journal of Philosophical Logic*, vol. 25, pp. 1–24.
4. Rao, C. R. 1973. *Linear Statistical Inference & Its Applications, 2nd Ed.*, Wiley, New York, NY.
5. Wilks, S. S. 1963. *Mathematical Statistics*, Wiley, New York, NY.
6. Goodman, I. R. and H. T. Nguyen 1999. "Probability Updating Using Second-Order Probabilities and Conditional Event Algebra," *Information Sciences*, vol. 121, pp. 295–347.
7. Bamber, D. 2000. "Entailment with Near Surety of Scaled Assertions of High Conditional Probability," *Journal of Philosophical Logic*, vol. 29, pp. 1–74.
8. Bamber, D. and I. R. Goodman. 2000. "New Uses of Second-Order Probability Techniques in Estimating Critical Probabilities in Command and Control Decision-Making," *Proceedings of the 2000 Command & Control Research & Technology Symposium*, Naval Postgraduate School, http://www.dodccrp.org/2000CCRTS/cd/html/pdf_papers/Track_4/124.pdf.
9. Dubois, D. and H. Prade. 1980. *Fuzzy Sets & Systems: Theory and Applications*, Academic Press, New York, NY.
10. Nguyen, H. T. and E. A. Walker. 1997. *A First Course in Fuzzy Logic*, CRC Press, New York, NY.
11. Goodman, I. R. 1998. "Random Sets and Fuzzy Sets: a Special Connection," *Proceedings of the International Conference on Multisource-Multisensor Information Fusion (Fusion'98)*, vol. 1, pp. 93–100.
12. Goodman, I. R. and H. T. Nguyen. 1985. *Uncertainty Models for Knowledge-Based Systems*, North-Holland Press, Amsterdam.
13. Goodman, I. R. and G. F. Kramer. 1997. "Extension of Relational and Conditional Event Algebra to Random Sets with Applications to Data Fusion," in *Random Sets: Theory & Applications* (J. Goutsias, R. P. Mahler, and H. T. Nguyen, eds.), Springer, New York, NY, pp. 209–242.
14. Goodman, I. R. and H. T. Nguyen. 1995. "Mathematical Foundations of Conditionals and Their Probabilistic Assignments," *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, vol. 3, no. 3 (September), pp. 247–339.
15. Goodman, I. R., R. P. Mahler, and H. T. Nguyen. 1997. *Mathematics of Data Fusion*, Kluwer Academic, Dordrecht, Holland.
16. Schweizer, B. and A. Sklar. 1983. *Probabilistic Metric Spaces*, North-Holland, Amsterdam.
17. Nelsen, R. B. 1999. *An Introduction to Copulas* (Lecture Notes in Statistics, no. 139), Springer, New York, NY.



I. R. Goodman

Ph.D. in Mathematics, Temple University, 1972

Current Research: Mathematical foundations of data fusion via conditional probabilistic logic; Boolean conditional event algebra; one-point random set representations of fuzzy logic.

18. Goodman, I. R. 1994. "A New Characterization of Fuzzy Logic Operators Producing Homomorphic-Like Relations with One Point Coverage of Random Sets," in *Advances in Fuzzy Theory & Technology*, (P. P. Wang, ed.), Duke University, Durham, NC, vol. 2, pp. 133–159.
19. Frank, M. J. 1979. "On the Simultaneous Associativity of F(x,y) and x+y-F(x,y)," *Aequationes Mathematicae*, vol. 19, pp. 194–226.
20. Zadeh, L. A.1985. "Syllogistic Reasoning as a Basis for Combination of Evidence in Expert Systems," *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-85)*, vol. 1, pp. 417–419.
21. Zadeh, L. A. 1978. "PRUF: a Meaning Representation Language for Natural Languages," *International Journal of Man–Machine Studies*, vol. 10, pp. 395–460.
22. Dubois, D. and H. Prade. 1988. "On Fuzzy Syllogisms," *Computational Intelligence*, vol. 4, pp. 171–179.
23. Kosko, B. 1992. *Neural Networks and Fuzzy Systems*, Prentice-Hall, Englewood Cliffs, NJ.

❖